



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------|-------------|----------------------|---------------------|------------------|
| 09/806,537 | 07/05/2001 | Y Tom Tang | PF-0612 USN | 1263 |

22428 7590 08/10/2004

FOLEY AND LARDNER
SUITE 500
3000 K STREET NW
WASHINGTON, DC 20007

EXAMINER

LANDSMAN, ROBERT S

ART UNIT PAPER NUMBER

1647

DATE MAILED: 08/10/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/806,537

Applicant(s)

TANG ET AL.

Examiner

Robert Landsman

Art Unit

1647

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 06 July 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) 7,8,16-18 and 20 is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-6,9-15 and 19 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 28 March 2001 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☒ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date 7/6/04.
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☒ Other: Sequence Comparisons A-E

Art Unit: 1647

DETAILED ACTION

1. Formal Matters

- A. The Election dated 7/6/04 has been entered into the record.
- B. Claims 1-20 were pending and were subject to restriction in the Office Action mailed 5/3/04. Applicants elected Group I, claims 1-6, 9-15 and 19, with traverse. The traversal will be discussed below. Claims 7, 8, 16-18 and 20 have been withdrawn as being drawn to a non-elected invention. Therefore, claims 1-6, 9-15 and 19 are the subject of this Office Action.

2. Answer to Traversal

- A. Applicants argue that SEQ ID NO:9, which is the polynucleotide encoding the polypeptide of SEQ ID NO:1 should be searched. This argument is persuasive. Therefore, SEQ ID NO:1 and 9 will be searched.

3. Specification

- A. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

The following title is suggested: G protein-coupled receptor proteins, encoding nucleic acids and methods of using the proteins.

4. Claim Objections

- A. Claims 1-6, 9-15 and 19 are objected to since they recite non-elected SEQ ID NOs.

5. Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

- A. Claims 1-6, 9-15 and 19 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by a specific, substantial and credible asserted utility or a well established utility. These claims are directed to a protein of SEQ ID NO:1, its encoding polynucleotide of SEQ ID NO:9 and methods of treating disorders. However, the invention encompassed by these claims has no apparent or disclosed patentable utility. This rejection is consistent with the current utility guidelines, published 1/5/01, 66 FR

Art Unit: 1647

1092. The instant application has provided a description of an isolated protein. However, the instant application does not disclose a specific and substantial biological role of this protein or its significance.

However, it is clear from the instant specification that the claimed receptor is what is termed an "orphan receptor" in the art. The instant application does not disclose the biological role of the claimed protein or its significance in relation to any disease state. Applicants disclose in the specification that the claimed receptor is homologous to human epididymis-specific seven transmembrane receptor HE6. However, homology alone is not sufficient. There is little doubt that, after complete characterization, this protein will probably be found to have a patentable utility. This further characterization, however, is part of the act of invention and, until it has been undertaken, Applicants' claimed invention is incomplete.

The instant situation is directly analogous to that of which was addressed in *Brenner v. Manson*, 148 U.S.P.Q. 689 (Sus. Ct, 1966), in which a novel compound which was structurally analogous to other compounds which were known to possess anticancer activity was alleged to be potentially useful as an antitumor agent in the absence of evidence supporting this utility. The court expressed the opinion that all chemical compounds are "useful" to the chemical arts when this term is given its broadest interpretation. However, the court held that this broad interpretation was not the intended definition of "useful" as it appears in 35 U.S.C. 101, which required that an invention must have either an immediate obvious or fully disclosed "real-world" utility. The court held that:

"The basic quid pro quo contemplated by the Constitution and the Congress for granting a patent monopoly is the benefit derived by the public from an invention with substantial utility," "[u]nless and until a process is refined and developed to this point - where specific benefit exists in currently available form - there is insufficient justification for permitting an applicant to engross what may prove to be a broad field," and "a patent is not a hunting license," "[i]t is not a reward for the search, but compensation for its successful conclusion."

The specification discloses that the polypeptide of the invention has sequence similarity to a known human epididymis-specific seven transmembrane receptor HE6. Based on the structural similarity, the specification asserts that the newly disclosed SEQ ID NO:1 and 9 have similar activities. The assertion that the disclosed proteins have biological activities similar to known receptors cannot be accepted in the absence of supporting evidence, because, generally, the art acknowledges that function cannot be predicted based solely on structural similarity to a protein found in the sequence databases. For example, Skolnick et al. (2000, Trends in Biotech. 18:34-39) state that knowing the protein structure by itself is insufficient to annotate a number of functional classes, and is also insufficient for annotating the specific details of protein function (see Box 2, p. 36). Similarly, Bork (2000, Genome Research 10:398-400) states that the error rate of functional annotations in the sequence database is considerable, making it

Art Unit: 1647

even more difficult to infer correct function from a structural comparison of a new sequence with a sequence database (see especially p. 399). Such concerns are also echoed by Doerks et al. (1998, Trends in Genetics 14:248-250) who state that (1) functional information is only partially annotated in the database, ignoring multi functionality, resulting in underpredictions of functionality of a new protein and (2) overpredictions of functionality occur because structural similarity often does not necessarily coincide with functional similarity. Smith et al. (1997, Nature Biotechnology 15:1222-1223) remark that there are numerous cases in which proteins having very different functions share structural similarity due to evolution from a common ancestral gene.

Brenner (1999, Trends in Genetics 15:132-133) argues that accurate inference of function from homology must be a difficult problem since, assuming there are only about 1000 major gene superfamilies in nature, then most homologs must have different molecular and cellular functions. Finally, Bork et al. (1996, Trends in Genetics 12:425-427) add that the software robots that assign functions to new proteins often assign a function to a whole new protein based on structural similarity of a small domain of the new protein to a small domain of a known protein. Such questionable interpretations are written into the sequence database and are then considered facts.

Therefore, based on the discussions above concerning the specific examples of structurally similar proteins that have different functions, along with the art's recognition that one cannot rely upon structural similarity alone to determine functionality, the specification fails to teach the skilled artisan the utility of the claimed polypeptide of SEQ ID NO:1 and the polynucleotides of SEQ ID NO:9 which are only known to be homologous to a known receptor. Therefore, the instant claims are drawn to a polynucleotide encoding a protein which has a yet undetermined function or biological significance. There is no actual and specific significance which can be attributed to said protein identified in the specification, or any association with a disease state. For this reason, the instant invention is incomplete. In the absence of a knowledge of the natural ligands or biological significance of this protein, there is no immediately obvious patentable use for it. To employ a protein of the instant invention in the identification of substances which bind to and/or mediate activity of the said receptor is clearly to use it as the object of further research which has been determined by the courts to be a non-patentable utility. Since the instant specification does not disclose a "real-world" use for said protein then the claimed invention is incomplete and, therefore, does not meet the requirements of 35 U.S.C. 101 as being useful.

Furthermore, since the nucleic acid and protein of the invention are not supported by a specific and substantial asserted utility or a well established utility, the vector, host cell as well as methods for producing and using the claimed polypeptide also lack utility.

Art Unit: 1647

6. Claim Rejections - 35 USC § 112, first paragraph - scope of enablement

The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

A. Claims 1-6, 9-15 and 19 are rejected under 35 U.S.C. 112, first paragraph, as failing to adequately teach how to use the instant invention. Specifically, since the claimed invention is not supported by a specific, substantial and credible asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

B. Furthermore, even if claim 1-6, 9-15 and 19 possessed utility under 35 USC 101, they would still be rejected under 35 U.S.C. 112, first paragraph, because the specification, while then being enabling for the protein of SEQ ID NO:1 and the polynucleotide of SEQ ID NO:9, does not reasonably provide enablement for proteins which are at least **"70% or 90%"** identical to SEQ ID NO:1 or 9, **"fragments"** thereof, or which **"hybridize"** to SEQ ID NO:9. Furthermore, the specification is not enabled for **"host cells"** which are not **"isolated,"** for **"pharmaceutical compositions"** or for **"methods of treating disorders"** with these compositions. The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make the invention commensurate in scope with these claims.

In In re Wands, 8USPQ2d, 1400 (CAFC 1988) page 1404, the factors to be considered in determining whether a disclosure would require undue experimentation include (1) the quantity of experimentation necessary, (2) the amount of direction or guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breadth of the claims.

First, the breadth of the claims is excessive with regard to claiming all proteins and polynucleotides which are "at least 70% or 90% identical" to SEQ ID NO:1 or 9, respectively, or those which "hybridize" under stringent conditions to SEQ ID NO:9. Nucleic acid molecules which are "at least 70% or 90% identical" to SEQ ID NO:9," which "hybridize" to SEQ ID NO:9, or which comprise fragments of SEQ ID NO: 9 would have one or more nucleic acid substitutions, deletions, insertions and/or additions to the polynucleotide of SEQ ID NO:9. Similarly, proteins which are "at least 70% or 90% identical" to the protein of SEQ ID NO:1, or comprise fragments thereof, would encode for a protein

Art Unit: 1647

with one or more amino acid substitutions, deletions, insertions and/or additions to the protein of SEQ ID NO:1.

Applicants provide no guidance or working examples of nucleic acid molecules which hybridize to SEQ ID NO:1, or of proteins which are at least 70% or 90% identical to SEQ ID NO:1, or polynucleotides which are at least 70% or 90% identical to SEQ ID NO:9, nor do they provide a *function* of these nucleic acid molecules, or proteins. Furthermore, fragments can be as few as one amino acid, or nucleic acid. Applicants have provided no guidance as to what critical residues are required to maintain the functional characteristics of any protein other than that of the full-length of SEQ ID NO:1 or the polynucleotide of SEQ ID NO:9. Furthermore, it is not predictable to one of ordinary skill in the art how to make a functional protein which is less than 100% identical to that of SEQ ID NO:1, or the polynucleotide of SEQ ID NO:9.

In addition, Applicants are not enabled for host cells which are not "isolated." As stand, claim 13 reads on gene-therapy since the cells transfected with the DNA could be in, for example, a human host. Claims 15 and 19 recite "pharmaceutical compositions." However, Applicants have not provided any guidance and working examples of the use of the protein of the invention to treat any disorder, nor are Applicants enabled for the scope of claim 19 since there are no examples of any disorders to be treated.

In summary, the breadth of the claims is excessive with regard to Applicants claiming all proteins and polynucleotides which are at least "70% or 90%" identical to SEQ ID NO:1 or 9, "fragments" thereof, or polynucleotides which "hybridize" to SEQ ID NO:9. There is also a lack of guidance and working examples of these nucleic acid molecules and proteins as well as which residues are critical for protein function. Similarly, the specification does not enable host cells which are not isolated, pharmaceutical compositions or methods of treating any claimed disorder. These factors, along with the lack of predictability to one of ordinary skill in the art as to how to make a functional proteins and polynucleotides other than that of SEQ ID NO:1 and 9, or how to use pharmaceutical compositions to treat the claimed disorders, leads the Examiner to hold that undue experimentation is necessary to practice the invention as claimed.

Art Unit: 1647

7. Claim Rejections - 35 USC § 112, first paragraph – written description

A. Claims 1-6, 9-15 and 19 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

These are genus claims. Nucleic acid molecules which are “**at least 70% or 90% identical**” to SEQ ID NO:9,” which “**hybridize**” to SEQ ID NO:9, or which comprise “**fragments**” of SEQ ID NO: 9 would have one or more nucleic acid substitutions, deletions, insertions and/or additions to the polynucleotide of SEQ ID NO:9. Similarly, proteins which are “at least 70% or 90% identical” to the protein of SEQ ID NO:1, or comprise fragments thereof, would encode for a protein with one or more amino acid substitutions, deletions, insertions and/or additions to the protein of SEQ ID NO:1.

The specification and claims do not indicate what distinguishing attributes are shared by the members of the genus. Thus the scope of the claims includes numerous structural variants, and the genus is highly variant because a significant number of structural differences between genus members is permitted. The specification and claims do not provide any guidance as to what changes should be made. Structural features that could distinguish compounds in the genus from others in the nucleic acid or protein class are missing from the disclosure. No common structural attributes identify the members of the genus. The general knowledge and level of skill in the art do not supplement the omitted description because specific, not general, guidance is what is needed. Since the disclosure fails to describe the common attributes or characteristics that identify members of the genus, and because the genus is highly variant, SEQ ID NO:1 and 9, or molecules which hybridize to the polynucleotides of SEQ ID NO:9 (which could be at least thousands of molecules) are insufficient to describe the genus. One of skill in the art would reasonable conclude that the disclosure fails to provide a representative number of species to describe the genus. Thus, Applicant was not in possession of the claimed genus at the time the invention was made.

This rejection also pertains to the claimed “**pharmaceutical compositions**” and “**methods of treating disorders.**” The specification does not provide adequate written description of “pharmaceutical compositions” to treat any disorder, nor does it provide written description of any disorders to be treated with these compositions.

Art Unit: 1647

8. Claim Rejections - 35 USC § 112, second paragraph

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

A. Claim 5 is vague and indefinite since the claim recites “stringent conditions.” It is not known what these conditions are. Nucleic acid molecules which hybridize under conditions of “low” stringency would not necessarily hybridize under conditions of “high” stringency. Furthermore, not all conditions of “high” or “low” stringency, for example, are the same. Therefore, it is required that Applicants amend the claims to recite the exact hybridization conditions without using indefinite phrases such as “*for example*” **without adding new matter.**

B. Claim 19 is rejected under 35 U.S.C. 112, second paragraph, as being incomplete for omitting essential steps, such omission amounting to a gap between the steps. See MPEP § 2172.01. The omitted steps are: a conclusion step demonstrating when the claimed method has been performed (i.e. when the objective of the claim has been reached). The claim recites a method of treating a disorder. However, there is no steps which determine, or identify, when the claimed disorder has been treated.

9. Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

A. Claims 1-6, 9-15 and 19 are rejected under 35 U.S.C. 102(a) as being anticipated by Wood et al. (WO 99/46281). The claims recite a polypeptide of SEQ ID NO:1, fragments thereof, as well as encoding polynucleotides, or fragments thereof. The claims also recite vectors, host cells, methods of making polypeptide, pharmaceutical compositions and methods of treating a disorder. Wood meet these limitations (Abstract; p296-299; Sequence Comparisons A, C, D). They teach a protein and polynucleotide which are 100% identical to SEQ ID NO:1 and 9 of the present invention. The artisan would immediately envision pharmaceutical compositions as well as methods of treating diseases (Examples 108 and 109 of page 306; lines 25-30 of page 132).

Art Unit: 1647

B. Claims 1-6 and 9-11 are rejected under 35 U.S.C. 102(a) as being anticipated by Liu et al. (Genomics 1999). Liu teach a polynucleotide encoding a protein (or fragment thereof) which is 99.9% identical to that of the present invention (Sequence Comparison B). This polynucleotide would be expected to hybridize to SEQ ID NO:9.

C. Claims 1-6, 9-11 are rejected under 35 U.S.C. 102(e) as being anticipated by US Patent 6,183,968. The patent discloses a polynucleotide which is 80% identical to that of the present invention (Sequence Comparison E). The artisan would immediately envision that this polynucleotide encodes a fragment of SEQ ID NO:1.

10. Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

A. Claims 12-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over either Liu or US Patent 6,183,968 each in view of Wood et al. The claims of the present invention as well as the teachings of all cited references are seen in the above rejection under 35 USC 102. Neither Liu, nor '968 teach vectors, host cells, or methods of making protein. However, Wood do teach these limitations.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have modified the invention of Wood et al. by substituting a cDNA in the polycloning region of the vector with the polynucleotide either Liu or '968 for the purpose of transfecting a host cell as taught by Wood. One of ordinary skill in the art would have been motivated to make this substitution in order to express the protein encoded by the introduced DNA in a host cell to perform ligand binding and functional assays. There would have been a reasonable expectation of success for a person of ordinary skill in the art to make this invention since these techniques are widely used in the art and are highly successful. The present invention, therefore, is *prima facie* obvious over the above references in the absence of evidence to the contrary.

Art Unit: 1647

11. Conclusion

A. No claim is allowable

Advisory information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Robert Landsman whose telephone number is (571) 272-0888. The examiner can normally be reached on Monday - Friday from 8:00 AM to 5:00 PM (Eastern time) and alternate Fridays from 8:00 AM to 5:00 PM (Eastern time).

If attempts to reach the examiner by telephone are unsuccessful, the Examiner's supervisor, Brenda Brumback, can be reached on (571) 272-0961.

Official papers filed by fax should be directed to (703) 872-9306. Fax draft or informal communications with the examiner should be directed to (571) 273-0888.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group receptionist whose telephone number is (571) 272-0700.

Robert Landsman, Ph.D.

Patent Examiner

Group 1600

July 16, 2004


ROBERT LANDSMAN
PATENT EXAMINER

Sequence Comparison

A

ID AAY41765 standard; protein; 693 AA.
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO1083 protein sequence.
 XX
 PN WO9946281-A2.
 PD 16-SEP-1999.
 PF 08-MAR-1999; 99WO-US005028.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 XX
 DR WPI; 1999-551358/46.
 DR N-PSDB; AAZ34292.
 XX
 SQ Sequence 693 AA;

Query Match 100.0%; Score 3604; DB 2; Length 693;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 693; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Qy | 1 | MTPQSLLOTTLFLLSLLFLVQGAHGRGHREDFRCSQRNQTHRSSLHYKPTPDLRISIEN | 60 |
| Db | 1 | MTPQSLLOTTLFLLSLLFLVQGAHGRGHREDFRCSQRNQTHRSSLHYKPTPDLRISIEN | 60 |
| Qy | 61 | SEEALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFFLLSDKASSLL | 120 |
| Db | 61 | SEEALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFFLLSDKASSLL | 120 |
| Qy | 121 | CFQHQQEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPHTAAHNASVDMCELK | 180 |
| Db | 121 | CFQHQQEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPHTAAHNASVDMCELK | 180 |
| Qy | 181 | RDLQLLSQFLKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKL | 240 |
| Db | 181 | RDLQLLSQFLKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKL | 240 |
| Qy | 241 | QPTAGLQDLHIHSRQEEEQSEIMEYSVLLPRTLFTQRTKGRSGEAEKRLLLVDFSSQALFQ | 300 |
| Db | 241 | QPTAGLQDLHIHSRQEEEQSEIMEYSVLLPRTLFTQRTKGRSGEAEKRLLLVDFSSQALFQ | 300 |
| Qy | 301 | DKNSSQVLGEKVLGIVVQNTKVANLTPVVLTFOHQLQPKNVTLCQVFWVEDPTLSSPGH | 360 |
| Db | 301 | DKNSSQVLGEKVLGIVVQNTKVANLTPVVLTFOHQLQPKNVTLCQVFWVEDPTLSSPGH | 360 |
| Qy | 361 | WSSAGCETVRRETQTSCFCNHLTYFAVLMVSSVEVDVAVHKHYLSLLSYVGCVVSAALCLV | 420 |
| Db | 361 | WSSAGCETVRRETQTSCFCNHLTYFAVLMVSSVEVDVAVHKHYLSLLSYVGCVVSAALCLV | 420 |
| Qy | 421 | TIAAYLCSRVPPLPCRRKPRDYTIKVHMNLLAVFLLDTSFLLSEPVALTGSEAGCRASAI | 480 |
| Db | 421 | TIAAYLCSRVPPLPCRRKPRDYTIKVHMNLLAVFLLDTSFLLSEPVALTGSEAGCRASAI | 480 |
| Qy | 481 | FLHFSLLTCLSWMGLEGYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVD | 540 |
| Db | 481 | FLHFSLLTCLSWMGLEGYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVD | 540 |
| Qy | 541 | NYGPIILAVHRTPEGVIYPSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLR | 600 |
| Db | 541 | NYGPIILAVHRTPEGVIYPSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLR | 600 |

Qy 601 PHTQKWSHVLTLGLSLVLGLPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWS 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 PHTQKWSHVLTLGLSLVLGLPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWS 660

Qy 661 MRLQARGGPSPLKSNSDSARLPISSGSTSSSRI 693
 ||||||||||||||||||||||||||||||||||||
 Db 661 MRLQARGGPSPLKSNSDSARLPISSGSTSSSRI 693

Sequence Comparison

AF106858
 LOCUS AF106858 2822 bp mRNA linear PRI 17-MAY-1999
 DEFINITION Homo sapiens G-protein-coupled receptor (GPR56) mRNA, complete cds.
 ACCESSION AF106858
 VERSION AF106858.1 GI:4836764
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2822)
 AUTHORS Liu, M., Parker, R.M., Darby, K., Eyre, H.J., Copeland, N.G.,
 Crawford, J., Gilbert, D.J., Sutherland, G.R., Jenkins, N.A. and
 Herzog, H.
 TITLE GPR56, a novel secretin-like human G-protein-coupled receptor gene
 JOURNAL Genomics 55 (3), 296-305 (1999)
 MEDLINE 99168899
 PUBMED 10049584
 REFERENCE 2 (bases 1 to 2822)
 AUTHORS Herzog, H.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-1999) Neurobiology, Garvan Institute, 384
 Victoria St., Sydney, NSW 2010, Australia

B

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 3.89e-271 | Length: | 2822 |
| Score: | 3599.00 | Matches: | 692 |
| Percent Similarity: | 99.86% | Conservative: | 0 |
| Best Local Similarity: | 99.86% | Mismatches: | 1 |
| Query Match: | 99.86% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-806-537A-1 (1-693) x AF106858 (1-2822)

Qy 1 MetThrProGlnSerLeuLeuGlnThrThrLeuPheLeuLeuSerLeuLeuPheLeuVal 20
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 163 ATGACTCCCCAGTCGCTGCTGCAGACGACACTGTTCTGCTGAGTCTGCTCTTCTGGTC 222

Qy 21 GlnGlyAlaHisGlyArgGlyHisArgGluAspPheArgPheCysSerGlnArgAsnGln 40
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 223 CAAGGTGCCACGGCAGGGGCCACAGGAAGACTTTCGCTTCTGCAGCCAGCGGAACCAG 282

Qy 41 ThrHisArgSerSerLeuHisTyrLysProThrProAspLeuArgIleSerIleGluAsn 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 283 ACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTGCGCATCTCCATCGAGAAC 342

Qy 61 SerGluGluAlaLeuThrValHisAlaProPheProAlaAlaHisProAlaSerArgSer 80
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 343 TCCGAAGAGGCCCTCACAGTCCATGCCCTTTCCCTGCAGCCACCCTGCTTCCCGATCC 402

Qy 81 PheProAspProArgGlyLeuTyrHisPheCysLeuTyrTrpAsnArgHisAlaGlyArg 100
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 403 TTCCCTGACCCAGGGGCTCTACCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGA 462

| | | | |
|----|------|--|------|
| Qy | 101 | LeuHisLeuLeuTyrGlyLysArgAspPheLeuLeuSerAspLysAlaSerSerLeu | 120 |
| Db | 463 | TTACATCTTCTCTATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTC | 522 |
| Qy | 121 | CysPheGlnHisGlnGluGluSerLeuAlaGlnGlyProProLeuLeuAlaThrSerVal | 140 |
| Db | 523 | TGCTTCCAGCACCAGGAGGAGAGCCTGGCTCAGGGCCCCCGCTGTTAGCCACTTCTGTC | 582 |
| Qy | 141 | ThrSerTrpTrpSerProGlnAsnIleSerLeuProSerAlaAlaSerPheThrPheSer | 160 |
| Db | 583 | ACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCAGTGCCGCCAGCTTCACCTTCTCC | 642 |
| Qy | 161 | PheHisSerProProHisThrAlaAlaHisAsnAlaSerValAspMetCysGluLeuLys | 180 |
| Db | 643 | TTCCACAGTCCTCCCCACAGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTCAAA | 702 |
| Qy | 181 | ArgAspLeuGlnLeuLeuSerGlnPheLeuLysHisProGlnLysAlaSerArgArgPro | 200 |
| Db | 703 | AGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCC | 762 |
| Qy | 201 | SerAlaAlaProAlaSerGlnGlnLeuGlnSerLeuGluSerLysLeuThrSerValArg | 220 |
| Db | 763 | TCGGCTGCCCCGCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGA | 822 |
| Qy | 221 | PheMetGlyAspMetValSerPheGluGluAspArgIleAsnAlaThrValTrpLysLeu | 240 |
| Db | 823 | TTCATGGGGGACATGGTGTCTTCGAGGAGGACCGGATCAACGCCACGGTATGGAAGCTC | 882 |
| Qy | 241 | GlnProThrAlaGlyLeuGlnAspLeuHisIleHisSerArgGlnGluGluGlnSer | 260 |
| Db | 883 | CAGCCACAGCCGCGCTCCAGGACCTGCACATCCACTCCCGGCAGGAGGAGGAGCAGAGC | 942 |
| Qy | 261 | GluIleMetGluTyrSerValLeuLeuProArgThrLeuPheGlnArgThrLysGlyArg | 280 |
| Db | 943 | GAGATCATGGAGTACTCGGTGTGCTGCCTCGAACACTCTTCAGAGGACGAAAGGCCGG | 1002 |
| Qy | 281 | SerGlyGluAlaGluLysArgLeuLeuLeuValAspPheSerSerGlnAlaLeuPheGln | 300 |
| Db | 1003 | AGCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAG | 1062 |
| Qy | 301 | AspLysAsnSerSerGlnValLeuGlyGluLysValLeuGlyIleValValGlnAsnThr | 320 |
| Db | 1063 | GACAAGAATTCCAGCCAAGTCTGGGTGAGAAAGTCTTGGGGATTGTGGTACAGAACACC | 1122 |
| Qy | 321 | LysValAlaAsnLeuThrGluProValValLeuThrPheGlnHisGlnLeuGlnProLys | 340 |
| Db | 1123 | AAAGTAGCCAACCTCACGGAGCCCGTGGTGCTCACTTTCAGCACCAGCTACAGCCGAAG | 1182 |
| Qy | 341 | AsnValThrLeuGlnCysValPheTrpValGluAspProThrLeuSerSerProGlyHis | 360 |
| Db | 1183 | AATGTGACTCTGCAATGTGTGTTCTGGGTGAAGACCCACATTGAGCAGCCCGGGGCAT | 1242 |
| Qy | 361 | TrpSerSerAlaGlyCysGluThrValArgArgGluThrGlnThrSerCysPheCysAsn | 380 |
| Db | 1243 | TGGAGCAGTGCTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGCAAC | 1302 |
| Qy | 381 | HisLeuThrTyrPheAlaValLeuMetValSerSerValGluValAspAlaValHisLys | 400 |
| Db | 1303 | CACTTGACCTACTTTGCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAG | 1362 |
| Qy | 401 | HisTyrLeuSerLeuLeuSerTyrValGlyCysValValSerAlaLeuAlaCysLeuVal | 420 |
| Db | 1363 | CACTACCTGAGCCTCCTCTCCTACGTGGGCTGTGTGCTCTCTGCCCTGGCCTGCCTTGTG | 1422 |

| | | | |
|----|------|--|------|
| Qy | 421 | ThrIleAlaAlaTyrLeuCysSerArgValProLeuProCysArgArgLysProArgAsp | 440 |
| Db | 1423 | ACCATTGCCGCCTACCTCTGCTCCAGGTGCCCTGCCGTGCAGGAGGAAACCTCGGGAC | 1482 |
| Qy | 441 | TyrThrIleLysValHisMetAsnLeuLeuLeuAlaValPheLeuLeuAspThrSerPhe | 460 |
| Db | 1483 | TACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTGGACACGAGCTTC | 1542 |
| Qy | 461 | LeuLeuSerGluProValAlaLeuThrGlySerGluAlaGlyCysArgAlaSerAlaIle | 480 |
| Db | 1543 | CTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATC | 1602 |
| Qy | 481 | PheLeuHisPheSerLeuLeuThrCysLeuSerTrpMetGlyLeuGluGlyTyrAsnLeu | 500 |
| Db | 1603 | TTCTCTGCACTTCTCCCTGCTCACCTGCCTTTCCTGGATGGGCCCTCGAGGGGTACAACCTC | 1662 |
| Qy | 501 | TyrArgLeuValValGluValPheGlyThrTyrValProGlyTyrLeuLeuLysLeuSer | 520 |
| Db | 1663 | TACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGC | 1722 |
| Qy | 521 | AlaMetGlyTrpGlyPheProIlePheLeuValThrLeuValAlaLeuValAspValAsp | 540 |
| Db | 1723 | GCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGAC | 1782 |
| Qy | 541 | AsnTyrGlyProIleIleLeuAlaValHisArgThrProGluGlyValIleTyrProSer | 560 |
| Db | 1783 | AACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCTTCC | 1842 |
| Qy | 561 | MetCysTrpIleArgAspSerLeuValSerTyrIleThrAsnLeuGlyLeuPheSerLeu | 580 |
| Db | 1843 | ATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGGCTCTTCAGCCTG | 1902 |
| Qy | 581 | ValPheLeuPheAsnMetAlaMetLeuAlaThrMetValValGlnIleLeuArgLeuArg | 600 |
| Db | 1903 | GTGTTTCTGTTC AACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGC | 1962 |
| Qy | 601 | ProHisThrGlnLysTrpSerHisValLeuThrLeuLeuGlyLeuSerLeuValLeuGly | 620 |
| Db | 1963 | CCCCACACCCAAAAGTGGTCACATGTGCTGACACTGCTGGGCCCTCAGCCTGGTCCTTGGC | 2022 |
| Qy | 621 | LeuProTrpAlaLeuIlePhePheSerPheAlaSerGlyThrPheGlnLeuValValLeu | 640 |
| Db | 2023 | CTGCCCTGGGCCTTGATCTTCTTCTCCTTTGCTTCTGGCACCTTCCAGCTTGTCGTCTC | 2082 |
| Qy | 641 | TyrLeuPheSerIleIleThrSerPheGlnGlyPheLeuIlePheIleTrpTyrTrpSer | 660 |
| Db | 2083 | TACCTTTTCAGCATCATCACCTCCTTCCAAGGCTTCCTCATCTTCATCTGGTACTGGTCC | 2142 |
| Qy | 661 | MetArgLeuGlnAlaArgGlyGlyProSerProLeuLysSerAsnSerAspSerAlaArg | 680 |
| Db | 2143 | ATGCGGCTGCAGGCCCGGGGTGGCCCTCCCCTCGAAGAGCAACTCAGACTGCGCCAGG | 2202 |
| Qy | 681 | LeuProIleSerSerGlySerThrSerSerSerArgIle | 693 |
| Db | 2203 | CTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATC | 2241 |

ID AAZ34292 standard; cDNA; 3819 BP.
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO1083 nucleotide sequence.
 XX
 PN WO9946281-A2.
 XX
 DR WPI; 1999-551358/46.
 DR P-PSDB; AAY41765.
 XX
 PS Claim 2; Fig 203; 530pp; English.
 XX
 SQ Sequence 3819 BP; 789 A; 1221 C; 996 G; 813 T; 0 U; 0 Other;

Sequence Comparison
 C

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 4.46e-314 | Length: | 3819 |
| Score: | 3604.00 | Matches: | 693 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-806-537A-1 (1-693) x AAZ34292 (1-3819)

| | | | |
|----|-----|--|-----|
| Qy | 1 | MetThrProGlnSerLeuLeuGlnThrThrLeuPheLeuLeuSerLeuLeuPheLeuVal | 20 |
| Db | 214 | ATGACTCCCCAGTCGCTGCTGCAGACGACACTGTTCTGCTGAGTCTGCTCTTCCTGGTC | 273 |
| Qy | 21 | GlnGlyAlaHisGlyArgGlyHisArgGluAspPheArgPheCysSerGlnArgAsnGln | 40 |
| Db | 274 | CAAGGTGCCACGGCAGGGGCCACAGGAAGACTTTCGCTTCTGCAGCCAGCGGAACCAG | 333 |
| Qy | 41 | ThrHisArgSerSerLeuHisTyrLysProThrProAspLeuArgIleSerIleGluAsn | 60 |
| Db | 334 | ACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTGCGCATCTCCATCGAGAAC | 393 |
| Qy | 61 | SerGluGluAlaLeuThrValHisAlaProPheProAlaAlaHisProAlaSerArgSer | 80 |
| Db | 394 | TCCGAAGAGGCCCTCACAGTCCATGCCCTTTCCTGCAGCCCACCCTGCTTCCCGATCC | 453 |
| Qy | 81 | PheProAspProArgGlyLeuTyrHisPheCysLeuTyrTrpAsnArgHisAlaGlyArg | 100 |
| Db | 454 | TTCCCTGACCCAGGGGCCTCTACCACTTCTGCCTCTACTGGAACCGACATGTGGGAGA | 513 |
| Qy | 101 | LeuHisLeuLeuTyrGlyLysArgAspPheLeuLeuSerAspLysAlaSerSerLeuLeu | 120 |
| Db | 514 | TTACATCTTCTCTATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTC | 573 |
| Qy | 121 | CysPheGlnHisGlnGluGluSerLeuAlaGlnGlyProProLeuLeuAlaThrSerVal | 140 |
| Db | 574 | TGCTTCCAGCACCAGGAGGAGCCTGGCTCAGGGCCCCCGCTGTTAGCCACTTCTGTC | 633 |
| Qy | 141 | ThrSerTrpTrpSerProGlnAsnIleSerLeuProSerAlaAlaSerPheThrPheSer | 160 |
| Db | 634 | ACCTCTGGTGAGCCCTCAGAACATCAGCCTGCCAGTGCCGCCAGCTTCACCTTCTCC | 693 |
| Qy | 161 | PheHisSerProProHisThrAlaAlaHisAsnAlaSerValAspMetCysGluLeuLys | 180 |
| Db | 694 | TTCCACAGTCCTCCCCACAGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTCAA | 753 |
| Qy | 181 | ArgAspLeuGlnLeuLeuSerGlnPheLeuLysHisProGlnLysAlaSerArgArgPro | 200 |
| Db | 754 | AGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCC | 813 |

| | | | |
|----|------|---|------|
| Qy | 201 | SerAlaAlaProAlaSerGlnGlnLeuGlnSerLeuGluSerLysLeuThrSerValArg | 220 |
| Db | 814 | TCGGCTGCCCGCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAACTGACCTCTGTGAGA | 873 |
| Qy | 221 | PheMetGlyAspMetValSerPheGluGluAspArgIleAsnAlaThrValTrpLysLeu | 240 |
| Db | 874 | TTCATGGGGGACATGGTGTCTTCGAGGAGGACCGGATCAACGCCACGGTGTGGAAGCTC | 933 |
| Qy | 241 | GlnProThrAlaGlyLeuGlnAspLeuHisIleHisSerArgGlnGluGluGlnSer | 260 |
| Db | 934 | CAGCCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGGCAGGAGGAGGAGCAGAGC | 993 |
| Qy | 261 | GluIleMetGluTyrSerValLeuLeuProArgThrLeuPheGlnArgThrLysGlyArg | 280 |
| Db | 994 | GAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCAGAGGACGAAAGGCCGG | 1053 |
| Qy | 281 | SerGlyGluAlaGluLysArgLeuLeuLeuValAspPheSerSerGlnAlaLeuPheGln | 300 |
| Db | 1054 | AGCGGGGAGGCTGAGAAGAGACTCCTCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAG | 1113 |
| Qy | 301 | AspLysAsnSerSerGlnValLeuGlyGluLysValLeuGlyIleValValGlnAsnThr | 320 |
| Db | 1114 | GACAAGAATTCCAGCCAAGTCCTGGGTGAGAAGTCTTGGGGATTGTGGTACAGAACACC | 1173 |
| Qy | 321 | LysValAlaAsnLeuThrGluProValValLeuThrPheGlnHisGlnLeuGlnProLys | 340 |
| Db | 1174 | AAAGTAGCCAACCTCACGGAGCCCGTGGTGCTCACTTTCCAGCACCAGCTACAGCCGAAG | 1233 |
| Qy | 341 | AsnValThrLeuGlnCysValPheTrpValGluAspProThrLeuSerSerProGlyHis | 360 |
| Db | 1234 | AATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCACATTGAGCAGCCCGGGGCAT | 1293 |
| Qy | 361 | TrpSerSerAlaGlyCysGluThrValArgArgGluThrGlnThrSerCysPheCysAsn | 380 |
| Db | 1294 | TGGAGCAGTGCTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGCAAC | 1353 |
| Qy | 381 | HisLeuThrTyrPheAlaValLeuMetValSerSerValGluValAspAlaValHisLys | 400 |
| Db | 1354 | CACCTTGACCTACTTTGCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAG | 1413 |
| Qy | 401 | HisTyrLeuSerLeuLeuSerTyrValGlyCysValValSerAlaLeuAlaCysLeuVal | 420 |
| Db | 1414 | CACTACCTGAGCCTCCTCTCTACGTGGGCTGTGTCTCTGCGCTGGCCTGCCTTGTC | 1473 |
| Qy | 421 | ThrIleAlaAlaTyrLeuCysSerArgValProLeuProCysArgArgLysProArgAsp | 440 |
| Db | 1474 | ACCATTGCCGCCTACCTCTGCTCCAGGGTGCCCTGCCGTGCAGGAGGAAACCTCGGGAC | 1533 |
| Qy | 441 | TyrThrIleLysValHisMetAsnLeuLeuLeuAlaValPheLeuLeuAspThrSerPhe | 460 |
| Db | 1534 | TACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTCTGCTGGACACGAGCTTC | 1593 |
| Qy | 461 | LeuLeuSerGluProValAlaLeuThrGlySerGluAlaGlyCysArgAlaSerAlaIle | 480 |
| Db | 1594 | CTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATC | 1653 |
| Qy | 481 | PheLeuHisPheSerLeuLeuThrCysLeuSerTrpMetGlyLeuGluGlyTyrAsnLeu | 500 |
| Db | 1654 | TTCTTGCACTTCTCCCTGCTCACCTGCCTTTCTGGATGGGCCTCGAGGGGTACAACCTC | 1713 |
| Qy | 501 | TyrArgLeuValValGluValPheGlyThrTyrValProGlyTyrLeuLeuLysLeuSer | 520 |
| Db | 1714 | TACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGCTACCTACTCAAGCTGAGC | 1773 |

| | | | |
|----|------|---|------|
| Qy | 239 | CAAACCCACACACAGACCTGCGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCA | 298 |
| Db | 357 | CAAACCCACACACAGACCTGCGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCA | 416 |
| Qy | 299 | TGCCCCCTTTCCCTGCAGCCCAACCTGCTTCCCGATCCTTCCCTGACCCAGGGGCCTCTA | 358 |
| Db | 417 | TGCCCCCTTTCCCTGCAGCCCAACCTGCTTCCCGATCCTTCCCTGACCCAGGGGCCTCTA | 476 |
| Qy | 359 | CCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTCTATGGCAAGCG | 418 |
| Db | 477 | CCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTCTATGGCAAGCG | 536 |
| Qy | 419 | TGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAG | 478 |
| Db | 537 | TGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAG | 596 |
| Qy | 479 | CCTGGCTCAGGGCCCCCGCTGTTAGCCACTTCTGTACCTCCTGGTGGAGCCCTCAGAA | 538 |
| Db | 597 | CCTGGCTCAGGGCCCCCGCTGTTAGCCACTTCTGTACCTCCTGGTGGAGCCCTCAGAA | 656 |
| Qy | 539 | CATCAGCCTGCCAGTGCCGCCAGCTTACCTTCTCCTTCCACAGTCTCCCCACACGGC | 598 |
| Db | 657 | CATCAGCCTGCCAGTGCCGCCAGCTTACCTTCTCCTTCCACAGTCTCCCCACACGGC | 716 |
| Qy | 599 | CGCTCACAAATGCCTCGGTGGACATGTGCGAGCTCAAAGGGACCTCCAGCTGCTCAGCCA | 658 |
| Db | 717 | CGCTCACAAATGCCTCGGTGGACATGTGCGAGCTCAAAGGGACCTCCAGCTGCTCAGCCA | 776 |
| Qy | 659 | GTTCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCGCCAGCCAGCA | 718 |
| Db | 777 | GTTCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCGCCAGCCAGCA | 836 |
| Qy | 719 | GTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTATGAGGGGACATGGTGTCTTT | 778 |
| Db | 837 | GTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTATGAGGGGACATGGTGTCTTT | 896 |
| Qy | 779 | CGAGGAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCACAGCCGGCCTCCAGGA | 838 |
| Db | 897 | CGAGGAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCACAGCCGGCCTCCAGGA | 956 |
| Qy | 839 | CCTGCACATCCACTCCCGGCAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCT | 898 |
| Db | 957 | CCTGCACATCCACTCCCGGCAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCT | 1016 |
| Qy | 899 | GCTGCCTCGAACACTCTTCCAGAGGACGAAAGGCCGGAGCGGGGAGGCTGAGAAGAGACT | 958 |
| Db | 1017 | GCTGCCTCGAACACTCTTCCAGAGGACGAAAGGCCGGAGCGGGGAGGCTGAGAAGAGACT | 1076 |
| Qy | 959 | CCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCCAGCCAAGTCCT | 1018 |
| Db | 1077 | CCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCCAGCCAAGTCCT | 1136 |
| Qy | 1019 | GGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCC | 1078 |
| Db | 1137 | GGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCC | 1196 |
| Qy | 1079 | CGTGGTGCTCACCTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTT | 1138 |
| Db | 1197 | CGTGGTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTT | 1256 |
| Qy | 1139 | CTGGGTTGAAGACCCACATTGAGCAGCCCGGGGCATTGGAGCAGTGCTGGGTGTGAGAC | 1198 |
| Db | 1257 | CTGGGTTGAAGACCCACATTGAGCAGCCCGGGGCATTGGAGCAGTGCTGGGTGTGAGAC | 1316 |

| | | | |
|----|------|---|------|
| Qy | 1199 | CGTCAGGAGAGAAAACCCAAACATCCTGCTTCTGCAACCACTTGACCTACTTTGCAGTGCT | 1258 |
| Db | 1317 | CGTCAGGAGAGAAAACCCAAACATCCTGCTTCTGCAACCACTTGACCTACTTTGCAGTGCT | 1376 |
| Qy | 1259 | GATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCCTCCTCTCCTA | 1318 |
| Db | 1377 | GATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCCTCCTCTCCTA | 1436 |
| Qy | 1319 | CGTGGGCTGTGTCTGCTCTCTGCCCTGGCCTGCCTTGTACCATTGCCGCCTACCTCTGCTC | 1378 |
| Db | 1437 | CGTGGGCTGTGTCTGCTCTCTGCCCTGGCCTGCCTTGTACCATTGCCGCCTACCTCTGCTC | 1496 |
| Qy | 1379 | CAGGGTGCCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAA | 1438 |
| Db | 1497 | CAGGGTGCCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAA | 1556 |
| Qy | 1439 | CCTGCTGCTGGCCGTCTTCCTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCT | 1498 |
| Db | 1557 | CCTGCTGCTGGCCGTCTTCCTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCT | 1616 |
| Qy | 1499 | GACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATCTTCCTGCACTTCTCCCTGCTCAC | 1558 |
| Db | 1617 | GACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATCTTCCTGCACTTCTCCCTGCTCAC | 1676 |
| Qy | 1559 | CTGCCTTTCTGGATGGGCCCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTT | 1618 |
| Db | 1677 | CTGCCTTTCTGGATGGGCCCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTT | 1736 |
| Qy | 1619 | TGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCAT | 1678 |
| Db | 1737 | TGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCAT | 1796 |
| Qy | 1679 | CTTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACATATGGCCCCATCATCTTGGC | 1738 |
| Db | 1797 | CTTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACATATGGCCCCATCATCTTGGC | 1856 |
| Qy | 1739 | TGTGCATAGGACTCCAGAGGGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCT | 1798 |
| Db | 1857 | TGTGCATAGGACTCCAGAGGGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCT | 1916 |
| Qy | 1799 | GGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTCTCTGTTCAACATGGCCAT | 1858 |
| Db | 1917 | GGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTCTCTGTTCAACATGGCCAT | 1976 |
| Qy | 1859 | GCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAAAAGTGGTCACA | 1918 |
| Db | 1977 | GCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAAAAGTGGTCACA | 2036 |
| Qy | 1919 | TGTGCTGACACTGCTGGGCCTCAGCCTGGTCCTTGGCCTGCCCTGGGCCTTGATCTTCTT | 1978 |
| Db | 2037 | TGTGCTGACACTGCTGGGCCTCAGCCTGGTCCTTGGCCTGCCCTGGGCCTTGATCTTCTT | 2096 |
| Qy | 1979 | CTCCTTTGCTTCTGGCACCTTCCAGCTTGTCTGCTCTACCTTTTCAGCATCATCACCTC | 2038 |
| Db | 2097 | CTCCTTTGCTTCTGGCACCTTCCAGCTTGTCTGCTCTACCTTTTCAGCATCATCACCTC | 2156 |
| Qy | 2039 | CTTCCAAGGCTTCCTCATCTTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGG | 2098 |
| Db | 2157 | CTTCCAAGGCTTCCTCATCTTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGG | 2216 |
| Qy | 2099 | CCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCCAGGCTCCCCATCAGCTCGGGCAGCAC | 2158 |
| Db | 2217 | CCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCCAGGCTCCCCATCAGCTCGGGCAGCAC | 2276 |

D

| | | | |
|----|------|---|------|
| Qy | 2159 | CTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCGG | 2218 |
| | | | |
| Db | 2277 | CTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCGG | 2336 |
| Qy | 2219 | CCTCGTCGCACACTGCCTGTGGCCCCGAGCCCGGCCAGCCCCAGGCCAGTCAGCCGCA | 2278 |
| | | | |
| Db | 2337 | CCTCGTCGCACACTGCCTGTGGCCCCGAGCCAGGCCAGCCCCAGGCCAGTCAGCCGCA | 2396 |
| Qy | 2279 | GACTTTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGACTCCC | 2338 |
| | | | |
| Db | 2397 | GACTTTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGACTCCC | 2456 |
| Qy | 2339 | GGGCTGGGCTTTTGAATTGGCCTTGGGGACTACTCGGCTCTCACTCAGCTCCACGGGAC | 2398 |
| | | | |
| Db | 2457 | GGGCTGGGCTTTTGAATTGGCCTTGGGGACTACTCGGCTCTCACTCAGCTCCACGGGAC | 2516 |
| Qy | 2399 | TCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTGTCCCCACATCTGTCCCAACCCAGCTG | 2458 |
| | | | |
| Db | 2517 | TCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTGTCCCCACATCTGTCCCAACCCAGCTG | 2576 |
| Qy | 2459 | GAGGCCTGGTCTCTCCTTATAACCCCTGGGCCCAGCCCTCATTGCTGGGGGCCAGGCCTT | 2518 |
| | | | |
| Db | 2577 | GAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATTGCTGGGGGCCAGGCCTT | 2636 |
| Qy | 2519 | GGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCCCTGCCTGGGACAGAAATGTGGC | 2578 |
| | | | |
| Db | 2637 | GGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCCCTGCCTGGGACAGAAATGTGGC | 2696 |
| Qy | 2579 | TCCAGTTGCTCTGTCTCTCGTGGTCACCCTGAGGGCACTCTGCATCCTCTGTCAATTTAA | 2638 |
| | | | |
| Db | 2697 | TCCAGTTGCTCTGTCTCTCGTGGTCACCCTGAGGGCACTCTGCATCCTCTGTCAATTTAA | 2756 |
| Qy | 2639 | CCTCAGGTGGCACCCAGGGCGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGG | 2698 |
| | | | |
| Db | 2757 | CCTCAGGTGGCACCCAGGGCGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGG | 2816 |
| Qy | 2699 | CGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGCAGCTCGCCTACCTCTGAGCCCAGGCC | 2758 |
| | | | |
| Db | 2817 | CGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGCAGCTCGCCTACCTCTGAGCCCAGGCC | 2876 |
| Qy | 2759 | CCCTCCCTCCCTCAGCCCCCAGTCCCTCCCTCCATCTTCCCTGGGGTTCTCCTCCTCTCC | 2818 |
| | | | |
| Db | 2877 | CCCTCCCTCCCTCAGCCCCCAGTCCCTCCCTCCATCTTCCCTGGGGTTCTCCTCCTCTCC | 2936 |
| Qy | 2819 | CAGGCCTCCTTGCTCCTTCGTTACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTG | 2878 |
| | | | |
| Db | 2937 | CAGGCCTCCTTGCTCCTTCGTTACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTG | 2996 |
| Qy | 2879 | GGGAGTGGTTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCC | 2938 |
| | | | |
| Db | 2997 | GGGAGTGGTTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCC | 3056 |
| Qy | 2939 | TCGGCCTGCCCCTGAGCCAGGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTC | 2998 |
| | | | |
| Db | 3057 | TCGGCCTGCCCCTGAGCCAGGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTC | 3116 |
| Qy | 2999 | CATCTGGGCCTTTGTATGAGCTGCATTGCCCTTGCTCACCCCTGACCAAGCACACGCCTCA | 3058 |
| | | | |
| Db | 3117 | CATCTGGGCCTTTGTATGAGCTGCATTGCCCTTGCTCACCCCTGACCAAGCACACGCCTCA | 3176 |
| Qy | 3059 | GAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGACCATGCCAGTC | 3118 |
| | | | |
| Db | 3177 | GAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGACCATGCCAGTC | 3236 |

Qy 3119 CCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGG 3178
 |||||
 Db 3237 CCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGG 3296
 |||||
 Qy 3179 CCTAGAGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCAAATAGCTCCAGGCGCC 3238
 |||||
 Db 3297 CCTAGAGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCAAATAGCTCCAGGCGCC 3356
 |||||
 Qy 3239 CTCGGCCGCCCATCATGGTTAATTCTGTCCAACAAACACACACGGGTAGATTGCTGGCCT 3298
 |||||
 Db 3357 CTCGGCCGCCCATCATGGTTAATTCTGTCCAACAAACACACACGGGTAGATTGCTGGCCT 3416
 |||||
 Qy 3299 GTTGTAGGTGGTAGGGACACAGATGACCGACCTGGTCACTCCTCCTGCCAACATTTCAGTC 3358
 |||||
 Db 3417 GTTGTAGGTGGTAGGGACACAGATGACCGACCTGGTCACTCCTCCTGCCAACATTTCAGTC 3476
 |||||
 Qy 3359 TGGTATGTGAGGCGTGCCTGAAGCAAGAACTCCTGGAGCTACAGGGACAGGGAGCCATCA 3418
 |||||
 Db 3477 TGGTATGTGAGGCGTGCCTGAAGCAAGAACTCCTGGAGCTACAGGGACAGGGAGCCATCA 3536
 |||||
 Qy 3419 TTCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTTCAATCTTGACCTTGA 3478
 |||||
 Db 3537 TTCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTTCAATCTTGACCTTGA 3596
 |||||
 Qy 3479 AGATGGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAG 3538
 |||||
 Db 3597 AGATGGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAG 3656
 |||||
 Qy 3539 TACATGTTTCATTGTAGAGAATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAA 3598
 |||||
 Db 3657 TACATGTTTCATTGTAGAGAATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAA 3716
 |||||
 Qy 3599 TCAGCTGTTGTAATCACCTAGCAAAAAAAAAAAAA 3632
 |||||
 Db 3717 TCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAAA 3750
 |||||

; Sequence 78, Application US/09276531
 ; Patent No. 6183968
 ; GENERAL INFORMATION:
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/276,531
 ; FILING DATE: Herewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/079,677
 ; FILING DATE: March 27, 1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lynn E. Murry, Ph.D.
 ; REGISTRATION NUMBER: 42,918
 ; REFERENCE/DOCKET NUMBER: PA-0008 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 78:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3090 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:

Sequence Comparison

E

| | | | |
|----|------|---|------|
| Qy | 601 | CTCACAATGCCTCGGTGGACATGTGCGAGCTCAAAGGGACCTCCAGCTGCTCAGCCAGT | 660 |
| Db | 36 | CGCACAAATNCCTCGGTGGACATGTGCGAGCTCAAAGGGACCTCCNGCTGCTCAGCCAGT | 95 |
| Qy | 661 | TCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCGCCAGCCAGCAGT | 720 |
| Db | 96 | TCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCGCCAGCCAGCAGT | 155 |
| Qy | 721 | TGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTCATGGGGGACATGGTGTCTTCG | 780 |
| Db | 156 | TGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTCATGGGGGACATGGTGTCTTCG | 215 |
| Qy | 781 | AGGAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCCAGCCGGCCTCCAGGACC | 840 |
| Db | 216 | AGGAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCCAGCCGGCCTCCAGGACC | 275 |
| Qy | 841 | TGCACATCCACTCCCGGCAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGC | 900 |
| Db | 276 | TGCACATCCACTCCCGGCAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGC | 335 |
| Qy | 901 | TGCCTCGAACACTCTTCCAGAGGACGAAAGGCCGGAGCGGGGAGGCTGAGAAGAGACTCC | 960 |
| Db | 336 | TGCCTCGAACACTCTTCCAGAGGACGAAAGGCCGGAGCGGGGAGGCTGAGAAGAGACTCC | 395 |
| Qy | 961 | TCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCAGCCAAGTCCTGG | 1020 |
| Db | 396 | TCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCAGCCAAGTCCTGG | 455 |
| Qy | 1021 | GTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCCG | 1080 |
| Db | 456 | GTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCCG | 515 |
| Qy | 1081 | TGGTGCTCACCTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCT | 1140 |
| Db | 516 | TGGTGCTCACCTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCT | 575 |
| Qy | 1141 | GGGTTGAAGACCCACATTGAGCAGCCCGGGGCATTGGAGCAGTGTGGGTGTGAGACCG | 1200 |
| Db | 576 | GGGTTGAAGACCCACATTGAGCAGCCCGGGGCATTGGAGCAGTGTGGGTGTGAGACCG | 635 |
| Qy | 1201 | TCAGGAGAGAAACCCAAACATCCTGCTTCTGCAACCATTGACCTACTTTGCAGTGCTGA | 1260 |
| Db | 636 | TCAGGAGAGAAACCCAAACATCCTGCTTCTGCAACCATTGACCTACTTTGCAGTGCTGA | 695 |
| Qy | 1261 | TGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCCTCCTCTCCTACG | 1320 |
| Db | 696 | TGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCCTCCTCTCCTACG | 755 |
| Qy | 1321 | TGGGCTGTGTCTCTCTGCCCTGGCCTGCCTTGTCAACATTGCCGCCCTACCTCTGCTCCA | 1380 |
| Db | 756 | TGGGCTGTGTCTCTCTGCCCTGGCCTGCCTTGTCAACATTGCCGCCCTACCTCTGCTCCA | 815 |
| Qy | 1381 | GGGTGCCCTTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACC | 1440 |
| Db | 816 | GGGTGCCCTTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACC | 875 |

2

Qy 1441 TGCTGCTGGCCGTCCTTCCTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCTGA 1500
| | | | |
Db 876 TGCTGCTGGCCGTCCTTCCTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCTGA 935
| | | | |
Qy 1501 CAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATCTTCCTGCACCTCTCCCTGCTCACCT 1560
| | | | |
Db 936 CAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATCTTCCTGCACCTCTCCCTGCTCACCT 995
| | | | |
Qy 1561 GCCTTTCTCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTTTG 1620
| | | | |
Db 996 GCCTTTCTCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTTTG 1055
| | | | |
Qy 1621 GCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCT 1680
| | | | |
Db 1056 GCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCT 1115
| | | | |
Qy 1681 TTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACATATGGCCCCATCATCTTGGCTG 1740
| | | | |
Db 1116 TTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACATATGGCCCCATCATCTTGGCTG 1175
| | | | |
Qy 1741 TGCATAGGACTCCAGAGGGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCTGG 1800
| | | | |
Db 1176 TGCATAGGACTCCAGAGGGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCTGG 1235
| | | | |
Qy 1801 TCAGCTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTTCCTGTTCAACAT-GGCCATG 1859
| | | | |
Db 1236 TCAGCTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTTCCTGTTCAACATGGGCCATG 1295
| | | | |
Qy 1860 CTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAAAAGT-GGTCACA 1918
| | | | |
Db 1296 CTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAAAAGTGGGTGACA 1355
| | | | |
Qy 1919 TGTGCTGACACTGCTGGGCCTCAGCCTGGTCTTGGCCTGCCCT-GGGCTTGATCTTCT 1977
| | | | |
Db 1356 TGTGCTGACACTGCTGGGCCTCAGCCTGGTCTTGGCCTGCCCTGGGCCTTGATCTTCT 1415
| | | | |
Qy 1978 TCTCCTTTGCTTCTGGCACCTTCCAGCTTGTCGTCTCTACCTTTTCAGCATCATCACCT 2037
| | | | |
Db 1416 TCTCCTTTGCTTCTGGCACCTTCCAGCTTGTCGTCTCTACCTTTTCAGCATCATCACCT 1475
| | | | |
Qy 2038 CCTTCCAAGGCTTCCTCATCTTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCCGGGGTG 2097
| | | | |
Db 1476 CCTTCCAAGGCTTCCTCATCTTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCCGGGGTG 1535
| | | | |
Qy 2098 GCCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCCAGGCTCCCCATCAGCTCGGGCAGCA 2157
| | | | |
Db 1536 GCCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCCAGGCTCCCCATCAGCTCGGGCAGCA 1595
| | | | |
Qy 2158 CCTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCG 2217
| | | | |
Db 1596 CCTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCG 1655
| | | | |
Qy 2218 GCCTCGTCGCACACTGCCTGTGGCCCCCGAGCCCCGAGCCCCAGGCCAGTCAGCCGC 2277
| | | | |
Db 1656 GCCTCGTCGCACACTGCCTGTGGCCCCCGAGCCCCGAGCCCCAGGCCAGTCAGCCGC 1715
| | | | |
Qy 2278 AGACTTTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGAC-GGACTC 2336
| | | | |
Db 1716 AGACTTTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGGACTC 1775
| | | | |
Qy 2337 CCGGGCTGGGCTTTTGAATTGGCCTTGGGGACTACTCGGCTCTCACTCAGTCCCACGGG 2396
| | | | |
Db 1776 CCGGGCTGGGCTTTTGAATTGGCCTTGGGGACTACTCGGCTCTCACTCAGTCCCACGGG 1835
| | | | |

E

| | | | |
|----|------|--|------|
| Qy | 2397 | ACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTGTCCCCACATCTGTCCCAACCCAGC | 2456 |
| | | | |
| Db | 1836 | ACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTGTCCCCACATCTGTCCCAACCCAGC | 1895 |
| Qy | 2457 | TGGAGGCCTGGTCTCTCCTTATAACCCCT-GGGCCCAGCCCTCATTGCTGGGGGCCAGGC | 2515 |
| | | | |
| Db | 1896 | TGGAGGCCTGGTCTCTCCTTATAACCCCTGGGGCCCAGCCCTCATTGCTGGGGGCCAGGC | 1955 |
| Qy | 2516 | CTTGGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGT | 2575 |
| | | | |
| Db | 1956 | CTTGGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGT | 2015 |
| Qy | 2576 | GGCTCCAGTTGCTCTGTCTCTCGTGGTCACCCCTGAGGGCACTCTGCATCCTCTGTCAATTT | 2635 |
| | | | |
| Db | 2016 | GGCTCCAGTTGCTCTGTCTCTCGTGGTCACCCCTGAGGGCACTCTGCATCCTCTGTCAATTT | 2075 |
| Qy | 2636 | TAACTCAGGTGGCACCCAGGGCGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCC | 2695 |
| | | | |
| Db | 2076 | TAACTCAGGTGGCACCCAGGGCGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCC | 2135 |
| Qy | 2696 | TGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGCTCGCCTACCTCTGAGCCCAG | 2755 |
| | | | |
| Db | 2136 | TGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGCTCGCCTACCTCTGAGCCCAG | 2195 |
| Qy | 2756 | GGCCCTCCCTCCCTCAGCCCCCAGTCCTCCCTCCATCTTCCCTGGGGTTCTCCTCCTC | 2815 |
| | | | |
| Db | 2196 | GGCCCTCCCTCCCTCAGCCCCCAGTCCTCCCTCCATCTTCCCTGGGGTTCTCCTCCTC | 2255 |
| Qy | 2816 | TCCAGGGCCTCCTTGCTCCTTCGTTACAGCT-GGGGGTCCCCGATTCCAATGCTGTTT | 2874 |
| | | | |
| Db | 2256 | TCCAGGGCCTCCTTGCTCCTTCGTTACAGCTGGGGGTCCCCGATTCCAATGCTGTTT | 2315 |
| Qy | 2875 | TTTGGGGAGTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTGTCTACTGCACA | 2934 |
| | | | |
| Db | 2316 | TTTGGGGAGTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTGTCTACTGCACA | 2375 |
| Qy | 2935 | AGCCTCGGCCTGCCCCCT-GAGCCAGGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCT | 2993 |
| | | | |
| Db | 2376 | AGCCTCGGCCTGCCCCCTGGAGCCAGGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCT | 2435 |
| Qy | 2994 | CTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTGCTCACCCCTGACCAAGCACACG | 3053 |
| | | | |
| Db | 2436 | CTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTGCTCACCCCTGACCAAGCACACG | 2495 |
| Qy | 3054 | CCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTTGGCAAGAACTGTGGACCATGC | 3113 |
| | | | |
| Db | 2496 | CCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTTGGCAAGAACTGTGGACCATGC | 2555 |
| Qy | 3114 | CAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGAC | 3173 |
| | | | |
| Db | 2556 | CAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGAC | 2615 |
| Qy | 3174 | ACTGGCCTAGAGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCAAATAGCTCCAG | 3233 |
| | | | |
| Db | 2616 | ACTGGCCTAGGGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCAAATAGCTCCAG | 2675 |
| Qy | 3234 | GCGCCCTCGGCCGCCCATCATGGTTAATTCTGTCCAACAAACACACGGGTAGATTGCT | 3293 |
| | | | |
| Db | 2676 | GCGCCCTCGGCCGCCCATCATGGTTAATTCTGTCCAACAAACACACGGGTAGATTGCT | 2735 |
| Qy | 3294 | GGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTGGTCACTCCTCCTGCCAACATT | 3353 |
| | | | |
| Db | 2736 | GGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTGGTCACTCCTCCTGCCAACATT | 2795 |

[illegible]